A NECESSARY OPTIMALITY CONDITION FOR THE LINEAR-QUADRATIC DAE CONTROL PROBLEM

ANDRÉ BACKES *

Abstract. We consider the linear-quadratic optimal control problem for a controlled differential algebraic equation (DAE). Under minimal assumptions on the DAE concerning index and regularity it will be possible to prove that the sufficient optimality condition given in papers of G. Kurina and R. März is also a necessary condition. This condition includes the solution of an appropriate boundary value problem and, in the special case of an explicit ordinary differential equation, the condition is equal to the well known necessary and sufficient condition of the classical linear-quadratic optimal control problem.

Key words. Linear-quadratic optimal control problem, descriptor system, differential algebraic equation, necessary optimality condition

AMS subject classifications. 49K15, 49N10, 34H05

1. Introduction. This paper deals with the optimal control problem that consists of minimizing a quadratic cost function over control functions $u : [t_0, T] \to \mathbb{R}^k$ and solutions $x : [t_0, T] \to \mathbb{R}^m$ of a linear differential algebraic equation (DAE) of the form

$$A(Dx)' + Bx = Cu \tag{1.1}$$

together with the initial value condition

$$D(t_0)P[x(t_0) - x^0] = 0 \tag{1.2}$$

with fixed $x^0 \in \mathbb{R}^m$ and a projector $P \in \mathbb{R}^{m\times m}$.

The coefficients in (1.1) are supposed to be continuous matrix functions of suitable dimensions and the theory concerning the DAE in the special form (1.1) is covered for example in [Mä1], [Mä3], [Mä4], [BaMä].

The paper aims at proving a necessary optimality condition for this linear-quadratic optimal control problem. From [Mä2] we have a corresponding sufficient optimality condition for the optimal control problem in the case $D(t_0)P = D(t_0)$, which consists in a solution of an appropriate boundary value problem that offers an optimal control. This sufficient condition was already given in [KuMä] in the more general context when the state equation has a slightly different structure and acts in general Hilbert spaces.
In the special case of an explicit ordinary differential equation the condition is equal to the well known necessary and sufficient condition of the classical linear-quadratic optimal control problem. The solution of the mentioned boundary value problem includes the solution of an appropriate adjoint DAE, which is no problem in the classical case. The main problem when proving this sufficient condition to be also a necessary condition is the verification that this adjoint DAE is really solvable.

*Humboldt University, Institute of Mathematics, site: Rudower Chaussee 25, D-10099 Berlin, Germany, backes@mathematik.hu-berlin.de

In [Jo] the optimal control problem is discussed even for a nonlinear DAE and a necessary optimality condition using a very similar boundary value problem is given. But here the solvability of the adjoint equation is assumed and it is left open under what conditions concerning the optimization problem it is really given.
In [BeLa] the special case of a linear-quadratic problem with a time invariant linear DAE is investigated and the necessary condition of [Jo] is cited. But here, too, there is no hint under which conditions the solvability of the adjoint DAE is given.
While [Jo] and [BeLa] do not use an index concept for analyzing the DAE, in [Me] and [KuMe] the linear-quadratic problem is treated by the idea to transform the DAE resp. the control problem into a new problem where the DAE has index 1. This approach gives the justification to think about the necessary optimality condition just in case of index-1 DAEs. However, as this approach transforms the original problem it does not provide any information about a necessary optimality condition using the original DAE resp. the adjoint DAE belonging to the original problem.
In [Ba] the special case of a linear-quadratic problem with an index-2 DAE is investigated. Here, a special cost function that leads to a causal adjoint DAE is considered so that it is solvable and it is possible to prove a necessary optimality condition.
The aim of this paper is to consider the linear-quadratic problem with more general linear DAEs without making any assumptions concerning the quadratic cost function.
In the following Section 2 we give a more detailed characterization of the optimal control problem, in particular, the set of admissible controls. We look at the sufficient optimality condition in [Mä2] and prove that, in the case of index-1 DAEs, it is also a necessary condition. We present a counterexample that shows that the condition may fail to be a necessary condition in the case of a DAE without index 1.
In Section 3 we consider the DAE in extended Hessenberg-form that was already used in [Ba], and which will be useful as a technical tool in proofs.
In Section 4 we cite a result from [BaKuMä] concerning the index of the DAE in the boundary value problem occurring in the sufficient optimality condition.
In Section 5 we consider a new DAE that is generated by extending the controlled DAE in such a way that the new DAE has index 1. This will be possible under a special rank condition for the matrix C in (1.1) that controls the input of the control function to the DAE. We consider an appropriate optimal control problem for the new index-1 DAE and prove the necessary optimality condition for the original optimal control problem by using the necessary condition for the index-1 case. This includes that we can prove the solvability of the adjoint DAE belonging to the original problem. To be more precise it will be possible to prove the necessary condition for the optimal control problem with the DAE (1.1) and the initial value condition (1.2) using an arbitrary projector P. Actually, the sufficient condition given in [Mä2] is only valid for the case $D(t_0)P = D(t_0)$, in general.

2. The linear-quadratic optimal control problem for a controlled DAE.

On the time interval $[t_0, T]$ we consider the linear controlled DAE

$$A(Dx)' + Bx = Cu \tag{2.1}$$

together with the initial value condition

$$D(t_0)P[x(t_0) - x^0] = 0 \tag{2.2}$$

with fixed $x^0 \in \mathbb{R}^m$ and a projector $P \in \mathbb{R}^{m \times m}$. Let $Q := I - P$ be the conjugated projector to P. The coefficients of the controlled DAE are supposed to be continuous matrix functions $A \in C([t_0,T],\mathbb{R}^{m\times n})$, $D \in C([t_0,T],\mathbb{R}^{n\times m})$, $B \in C([t_0,T],\mathbb{R}^{m\times m})$ and $C \in C([t_0,T],\mathbb{R}^{m\times k})$.

For the DAE (2.1) we assume a properly stated leading term that means,

$$\ker A(t) \oplus \operatorname{im} D(t) = \mathbb{R}^n, \quad t \in [t_0,T], \tag{2.3}$$

and there exists a continuous differentiable projector function
$R \in C^1([t_0,T],\mathbb{R}^{n\times n})$ with $\operatorname{im} R(t) = \operatorname{im} D(t)$ and $\ker\, R(t) = \ker\, A(t)$ for $t \in [t_0,T]$. The matrix $G_0 := AD$ has constant rank on the interval $[t_0,T]$.

We use here the matrices, sets, and projectors of the matrix sequence defined in [Mä1]. Among others, we have

$$\begin{aligned}
&N_0 := \ker G_0 = \ker D\\
&Q_0 \in \mathbb{R}^{m\times m}, \quad Q_0^2 = Q_0, \quad \operatorname{im} Q_0 = N_0\\
&P_0 := I - Q_0\\
&G_1 := G_0 + BQ_0 = AD + BQ_0\\
&N_1 := \ker G_1\\
&Q_1 \in \mathbb{R}^{m\times m}, \quad Q_1^2 = Q_1, \quad \operatorname{im} Q_1 = N_1\\
&P_1 := I - Q_1.
\end{aligned} \tag{2.4}$$

The definition of these matrices and sets is meant pointwise for each $t \in [t_0,T]$.

<u>The regular DAE with index 1</u>

With the help of the matrices defined in (2.4) we can define what we understand when speaking of a regular DAE with index 1 (cf. [Mä1]):

Definition 2.1. *The DAE (2.1) is called regular with index 1 if the matrix G_1 is nonsingular.*

In [Mä1] the matrix sequence (2.4) is continued and it is defined what is meant by a regular DAE with index $\mu \in \mathbb{N}$. In this paper we just consider the case $\mu = 1$, or the other possibility that we do not assume anything concerning the regularity and the index of the DAE.

For an optimality problem the relation between the control function u and a corresponding solution x of (2.1), (2.2) is important. If the DAE is regular with index 1, then we have the following nice Existence Theorem for the solution of the associated initial value problem (e.g. [Mä3]):

Theorem 2.2. *(Solution of the DAE in case of index 1)*
We consider the initial value problem

$$\begin{cases} A(Dx)' + Bx = q \\ D(t_0)P_0(t_0)(x(t_0) - x^0) = 0, \quad x^0 \in \mathbb{R}^m, \end{cases} \tag{2.5}$$

for a regular DAE with index 1. Then, for every function $q \in C([t_0,T],\mathbb{R}^m)$, there exists a unique solution $x \in C^1_D([t_0,T],\mathbb{R}^m)$ of (2.5) and the estimation

$$\|x\|_\infty \le L\left(\|D(t_0)x^0\| + \|q\|_\infty\right) \tag{2.6}$$

is valid with a constant L.

The optimal control problem for the general DAE

As we want to consider only weak assumptions on the regularity and the index of the controlled DAE, we define, with regard to the examination of a cost function and a corresponding optimization problem, what we want to understand when speaking of the set of admissible controls:

Definition 2.3. *(admissible control and admissible pair)*
A control function $u \in C([t_0,T],\mathbb{R}^k)$ is called admissible if there exists a corresponding solution $x \in C^1_D([t_0,T],\mathbb{R}^m)$ of (2.1), (2.2). By

$$\mathcal{Z} \subset C([t_0,T],\mathbb{R}^k) \tag{2.7}$$

we denote the set of all admissible control functions.
For a control function $u \in C([t_0,T],\mathbb{R}^k)$,

$$\mathcal{L}_u \subset C^1_D([t_0,T],\mathbb{R}^m) \tag{2.8}$$

denotes the set of all corresponding solutions of the initial value problem (2.1), (2.2). A pair (x,u) with $u \in \mathcal{Z}$ and $x \in \mathcal{L}_u$ we call an admissible pair.

Remark 2.4. *Obviously it holds that*

$$\mathcal{Z} = \left\{ u \in C([t_0,T],\mathbb{R}^k) \quad | \quad \mathcal{L}_u \neq \emptyset \right\}. \tag{2.9}$$

Example 2.5. *(Sets of admissible controls and admissible pairs)*
For some well-known examples we consider what the concept of admissible control resp. admissible pair really means:

1. *Regular DAE with index 1, $P = P_0(t_0)$*

According to Theorem 2.2 it holds here that $\mathcal{Z} = C([t_0,T],\mathbb{R}^k)$ and for every $u \in \mathcal{Z}$ the set $\mathcal{L}_u$ contains exactly one element.

2. *Regular DAE with index 2, $P = P_0(t_0)P_1(t_0)$*

According to the appropriate theorem for the index 2 case (e.g., [Mä3]) we have $\mathcal{Z} = C^1_{DQ_1G_2^{-1}C}([t_0,T],\mathbb{R}^k)$ here and for every $u \in \mathcal{Z}$ the set $\mathcal{L}_u$ contains exactly one element.

3. *Regular DAE with index 2, $P = P_0(t_0)$*

Here we have

$$\begin{aligned} \mathcal{Z} \quad = \quad & \Big\{ u \in C^1_{DQ_1G_2^{-1}C}([t_0,T],\mathbb{R}^k) \quad | \\ & D(t_0)Q_1(t_0)G_2^{-1}(t_0)C(t_0)u(t_0) = D(t_0)Q_1(t_0)x^0 \Big\} \end{aligned} \tag{2.10}$$

and for every $u \in \mathcal{Z}$ the set $\mathcal{L}_u$ contains exactly one element.

The first and the second example show that we can obtain a somehow nice structure of the set $\mathcal{Z}$ if we choose the projector P in such a way that the initial value condition (2.2) and the index of the DAE fit together such that we have a nice Existence Theorem for the solution of the initial value problem (2.1), (2.2). In [Mä4] it is shown that such a suitable projector can always be found if the DAE is regular with some index μ.
The latter example shows that the set $\mathcal{Z}$ may really have a complex structure if, for a regular DAE with index μ, the projector P is not chosen in this way. An even more complicated structure of the set $\mathcal{Z}$ can be expected if we consider a nonregular DAE without any index. $\square$

Now let us consider a quadratic cost function J in the form

$$J(x,u) = \frac{1}{2}x^T(T)Vx(T) + \frac{1}{2}\int_{t_0}^{T} \left\{x^T W x + 2x^T S u + u^T K u\right\} \, dt \qquad (2.11)$$

that has to be minimized on the set of admissible pairs (x,u).

Therefore we consider the matrix $V \in \mathbb{R}^{m\times m}$ and the continuous matrix functions $W \in C([t_0,T],\mathbb{R}^{m\times m})$, $S \in C([t_0,T],\mathbb{R}^{m\times k})$ and $K \in C([t_0,T],\mathbb{R}^{k\times k})$ with $V^T = V$, $W^T(t) = W(t)$ and $K^T(t) = K(t)$ for all $t \in [t_0,T]$.

The matrices V and $\begin{pmatrix} W(t) & S(t) \\ S^T(t) & K(t) \end{pmatrix}$, $t \in [t_0,T]$, are considered to be positive semidefinite.

These standard assumptions correspond to the assumptions of the classical linear quadratic optimal control problem for an explicit ordinary differential equation. For our DAE control problem we make the additional assumption concerning the matrix V that

$$Vz = 0 \quad \text{for every} \quad z \in N_0(T) = \ker D(T), \qquad (2.12)$$

which is trivially fulfilled in the case of a nonsingular matrix D. Thus, for the classical problem, condition (2.12) is in fact no additional assumption.
Later we will see that the property (2.12) makes sense as a condition that assures the solvability of the final value problem for the adjoint DAE. [KuMä] contains already an example of the optimal control problem showing that the condition to V makes sense here.

So, altogether, we consider the optimization problem

$$J(x,u) \to \text{Min} \quad , \quad (x,u) \quad \text{admissible pair for} \quad (2.1),(2.2). \qquad (2.13)$$

Here we consider the case of a fixed time interval $[t_0,T]$, this means, the final time T is not a value that has to be optimized. Moreover, the range of values for the control function is not restricted, this means, we just consider the case that $u(t) \in \mathcal{U} = \mathbb{R}^k$.

Optimality conditions

With regard to the analysis of the optimization problem and the specification of optimality conditions we define

Definition 2.6. *(admissible variation)*
Consider an admissible pair (x, u) *, then* $(\delta u, \delta x)$ *is called an* *admissible variation* *of* (x, u) *if* $\delta u \in C([t_0, T], \mathbb{R}^k)$ *and* $\delta x \in C^1_D([t_0, T], \mathbb{R}^m)$ *are such that* $(u + \delta u, x + \delta x)$ *is again an admissible pair.*

Definition 2.7. *(optimal pair)*
An admissible pair (x_*, u_*) *is called* *optimal* *if for every admissible variation* $(\delta x, \delta u)$ *of* (x_*, u_*) *the property*

$$J(x_* + \delta x, u_* + \delta u) \geq J(x_*, u_*) \tag{2.14}$$

is valid.

Definition 2.8. *(optimal control and optimal trajectory)*
An admissible control u_* *is called an* *optimal control* *if there exists an* $x_* \in \mathcal{L}_{u_*}$ *such that* (x_*, u_*) *is an optimal pair. Every* x_* *with this property is called an* *optimal trajectory.*

If we consider the optimization problem (2.13) together with the special projector $P = P(t_0)$ in the initial value condition for the controlled DAE, then we have the following sufficient optimality condition from [Mä2]:

Theorem 2.9. *(sufficient optimality condition)*
Consider $x_* \in C^1_D([t_0, T], \mathbb{R}^m)$, $\lambda_* \in C^1_{A^T}([t_0, T], \mathbb{R}^m)$ *and* $u_* \in C([t_0, T], \mathbb{R}^k)$ *such that the triple* (x_*, λ_*, u_*) *is a solution of the boundary value problem*

$$(BVP)\left\{\begin{array}{rcl} A(Dx)' &=& -Bx + Cu \\ D^T(A^T\lambda)' &=& Wx + B^T\lambda + Su \\ 0 &=& S^Tx - C^T\lambda + Ku \\ && \\ D(t_0)P_0(t_0)(x(t_0) - x^0) &=& 0 \\ D^T(T)A^T(T)\lambda(T) &=& -Vx(T). \end{array}\right. \tag{2.15}$$

Then (x_*, u_*) *is an optimal pair, i.e.,* u_* *is an optimal control and* x_* *is a corresponding optimal trajectory.*

Remark 2.10. *In general the sufficient optimality condition is just valid for the optimization problem with the a projector* P *that satisfies* $D(t_0)P = D(t_0)$*. This is fulfilled for the projector* $P = P_0(t_0)$ *because of the property* $D(t_0)Q_0(t_0) = 0$*.*
However, for our task to obtain a necessary optimality condition we want to consider the more general optimization problem with an arbitrary projector P *in the initial value condition for the controlled DAE.*

The question arises whether the sufficient condition is also a necessary condition in general. This is important to know if we want to compute an optimal control by solving the boundary value problem (2.15). For example, we can think of a numerical method for computing an optimal control that is based on the idea to solve (2.15). If the condition is not necessary, then it might happen that there exists an optimal

control u_* for (2.13) but (2.15) does not have a corresponding solution (x_*, λ_*, u_*). Then it would be impossible to compute the optimal control u_* in this way.
A simple counterexample shows that it can really happen that the sufficient condition is not necessary:

Example 2.11. *(the sufficient optimality condition is not necessary)*
We consider the initial value problem

$$\left\{\begin{array}{rcll} \dot{x}_1 & = & u \quad , & x_1(0) = x_1^0 \\ \dot{x}_2 + x_3 & = & u \quad , & x_2(0) = 0 \\ x_2 & = & 0 & \end{array}\right. \tag{2.16}$$

for $x(t) = (x_1(t), x_2(t), x_3(t)) \in \mathbb{R}^3, \ t \in [t_0, T], \ x_1^0 \in \mathbb{R}.$

The controlled DAE (2.16) is regular with index 2. For every function $u \in \mathcal{Z} = C([t_0, T], \mathbb{R}^k)$ *we have a unique solution*

$$\left\{\begin{array}{rcll} x_1(t) & = & x_1^0 + \int_0^t u(s)\, ds & \\ x_2(t) & = & 0 & \\ x_3(t) & = & u(t) \quad , & t \in [t_0, T]. \end{array}\right. \tag{2.17}$$

We consider the cost function

$$\begin{aligned} J(x,u) = J(u) \quad &= \quad \tfrac{1}{2}x_1^2(T) + \tfrac{1}{2}\int_0^T x_3^2 + u^2\, dt \\ &= \quad \tfrac{1}{2}x_1^2(T) + \tfrac{1}{2}\int_0^T 2u^2\, dt, \end{aligned} \tag{2.18}$$

which has the special property to be also a cost function for the controlled explicit ordinary differential equation

$$\dot{x}_1 = u \quad , \quad x_1(0) = x_1^0. \tag{2.19}$$

Consider $\mu \in \mathbb{R}$ *to be the adjoint variable to* $x_1 \in \mathbb{R}$. *Then the final value problem for the adjoint equation belonging to the optimal control problem (2.19), (2.18) is*

$$\dot{\mu} = 0 \quad , \quad \mu(T) = -x_1(T), \tag{2.20}$$

and has the constant solution $\mu(t) = -x_1(T)$.

Thus, the well-known necessary condition for the classical linear-quadratic optimal control problem for explicit ordinary differential equations

$$\mu = 2u \quad \Longleftrightarrow \quad u = \frac{1}{2}\mu \tag{2.21}$$

provides the fact that the optimal control is a constant function.
For $u(t) = u_0 \in \mathbb{R}$ *we have*

$$J(u_0) = \frac{1}{2}(x_0 + u_0 T)^2 + u_0^2 T, \quad u_*(t) = -\frac{x_0}{2+T} \tag{2.22}$$

$$x_*(T) = \frac{2x_0}{2+T}, \quad J(u_*) = \frac{x_0^2}{2+T}. \tag{2.23}$$

Here the control function u_* is also an optimal control for the DAE-optimization-problem. The final value problem for the adjoint DAE from (2.15) is

$$\left\{\begin{array}{rcl} \dot{\lambda}_1 &=& 0 \quad , \quad \lambda_1(T) = -x_1(T) \\ \dot{\lambda}_2 - \lambda_3 &=& 0 \quad , \quad \lambda_2(T) = 0 \\ -\lambda_2 &=& x_3 \end{array}\right. \tag{2.24}$$

and it has for a control function $u \in C^1[t_0, T]$ with $u(T) = 0$ the unique solution

$$\left\{\begin{array}{rcl} \lambda_1(t) &=& -x_1^0 \\ \lambda_2(t) &=& -u(t) \\ \lambda_3(t) &=& -\dot{u}(t) \quad , \quad t \in [t_0, T]. \end{array}\right. \tag{2.25}$$

But for $x_0 \neq 0$ the optimal control u_* does not satisfy the condition $u_*(T) = 0$ here so that the final value problem for the adjoint DAE is not solvable. Thus, we have an optimal pair (x_*, u_*) and there is no λ_* such that (x_*, λ_*, u_*) is a solution of the boundary value problem. Obviously, the sufficient condition from Theorem 2.9 is not necessary here.
Although the optimal control is a very simple function in this example, it is not possible to compute the optimal control by solving the boundary value problem (2.15). □

The question whether the existence of a solution to a boundary value problem as in Theorem 2.9 is also a necessary condition for an optimal control was answered in the affirmative in [Ba] for the case of a regular DAE with index 2 together with the suitable initial value condition

$$D(t_0)P_1(t_0)(x(t_0) - x^0) = 0 \tag{2.26}$$

and the special cost function with the properties

$$\left\{\begin{array}{l} V(I - P_0(T)P_1(T)) = 0 \\ WT = 0 \\ S^T T = 0. \end{array}\right. \tag{2.27}$$

Here T is a projector onto the non-causal component of the state variable, which contains the derivative of the control function that occurs in the solution of the DAE in case of index 2. This component is <u>not</u> valuated by a cost function that has the properties (2.27).

<u>The necessary optimality condition in the index 1 case</u>

For the optimization problem (2.13) in case of a regular DAE with index 1 Theorem 2.2 justifies to speak of <u>the</u> solution x of (2.1), (2.2) and to write $J(u)$ instead of $J(x, u)$.
In case of a regular DAE with index 1 we can prove that the sufficient condition from Theorem 2.9 is also a necessary condition:

Theorem 2.12. *(Necessary optimality condition in case of index 1)*
Assume that the controlled DAE in the optimization problem (2.13) is regular with index 1 and in the initial condition (2.2) we have $P = P_0(t_0)$. Let $u_* \in C([t_0, T], \mathbb{R}^k)$

be an optimal control and $x_* \in C_D^1([t_0,T],\mathbb{R}^m)$ *the corresponding optimal trajectory. Then there exists a function*

$$\lambda_* \in C_{A^T}^1([t_0,T],\mathbb{R}^m)$$

such that (x_*,λ_*,u_*) *is a solution of the boundary value problem*

$$\left\{\begin{array}{rcl} A(Dx)' &=& -Bx+Cu \\ D^T(A^T\lambda)' &=& Wx+B^T\lambda+Su \\ 0 &=& S^Tx-C^T\lambda+Ku \\ \\ D(t_0)P_0(t_0)(x(t_0)-x^0) &=& 0 \\ D^T(T)A^T(T)\lambda(T) &=& -Vx(T). \end{array}\right. \tag{2.28}$$

Proof. Consider an arbitrary variation $\delta u \in C([t_0,T],\mathbb{R}^k)$ of the optimal control u_*. Denote by $\delta x \in C_D^1([t_0,T],\mathbb{R}^m)$ the corresponding variation of the optimal trajectory x_*. Then $(x_*+\delta x, u_*+\delta u)$ is an admissible variation of the optimal pair (x_*,u_*) with

$$D(t_0)\delta x(t_0) = 0 \tag{2.29}$$

and the property (2.6) provides the estimation

$$\|\delta x\|_\infty \le L\|\delta u\|_\infty \tag{2.30}$$

for some $L \ge 0$.

Because of the linearity of the DAE (2.1) also $(x_*+\epsilon\delta x, u_*+\epsilon\delta u)$ is an admissible pair for every $\epsilon \in [0,1]$. Thus, the directional derivative of J at u_* exists for the direction δu

$$\delta J(u_*)\delta u = \lim_{\epsilon\to 0} \frac{J(u_*+\epsilon\delta u)-J(u_*)}{\epsilon} \tag{2.31}$$

and has to vanish because of the property of u_* to be optimal. Thus, we have the necessary optimality condition

$$\delta J(u_*)\delta u = 0\,. \tag{2.32}$$

The Hamilton-function combines the cost function that has to be minimized with the DAE using the adjoint variable $\lambda \in \mathbb{R}^m$:

$$\begin{aligned} &H : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R} \to \mathbb{R}, \\ &H(x,\lambda,u,t) := -\tfrac{1}{2}x^TWx - x^TSu - \tfrac{1}{2}u^TKu + \lambda^T(-Bx+Cu). \end{aligned} \tag{2.33}$$

With the help of the Hamilton-function we can formulate the cost function J in the form

$$J(u) = \frac{1}{2}x^T(T)Vx(T) + \int_{t_0}^T \left\{\lambda^T A(Dx)' - H(x,\lambda,u,t)\right\} dt. \tag{2.34}$$

According to [BaMä] the adjoint DAE is also regular with index 1, hence we can choose $\lambda_* \in C_D^1([t_0,T],\mathbb{R}^m)$ as a solution to the final value problem

$$\left\{\begin{array}{l} D^T(A^T\lambda)' = -H_x^T(x_*,\lambda,u_*,t) = B^T\lambda + Wx_* + Su_* \\ D^T(T)A^T(T)\lambda(T) = -Vx_*(T). \end{array}\right. \tag{2.35}$$

In order to calculate the directional derivative of J we consider the difference $J(u_* + \delta u) - J(u_*)$, and using the property (2.30) the Taylor expansion yields

$$\begin{aligned}
\Delta J &= J(u_* + \delta u) - J(u_*) \\
&= \frac{1}{2}(x_*(T) + \delta x(T))^T V (x_*(T) + \delta x(T)) - \frac{1}{2} x_*^T(T) V x_*(T) \\
&\quad + \int_{t_0}^{T} \left\{ \lambda_*^T A(D(x_* + \delta x))' - \lambda_*^T A(Dx_*)' \right\} dt \\
&\quad - \int_{t_0}^{T} \{ H(x_* + \delta x, \lambda_*, u_* + \delta u, t) - H(x_*, \lambda_*, u_*, t) \} \, dt \\
&= x_*^T(T) V \delta x(T) + \int_{t_0}^{T} \left\{ \lambda_*^T A(D\delta x)' - H_x \delta x - H_u \delta u \right\} dt + o(\|\delta u\|_\infty).
\end{aligned}$$

Here we use the abbreviations

$$\begin{aligned}
H_x &:= H_x(x_*, \lambda_*, u_*, t) = -x_*^T W - u_*^T S - \lambda^T B \\
H_u &:= H_u(x_*, \lambda_*, u_*, t) = -x_*^T S + \lambda^T C - u_*^T K.
\end{aligned} \tag{2.36}$$

Partial integration yields

$$\int_{t_0}^{T} \lambda_*^T A(D\delta x)' \, dt = \left[\lambda_*^T A D \delta x \right]_{t_0}^{T} - \int_{t_0}^{T} (\lambda_*^T A)' D \delta x \, dt \tag{2.37}$$

and thus we have

$$\begin{aligned}
\Delta J \quad = \quad & x_*^T(T) V \delta x(T) + \left[\lambda_*^T A D \delta x \right]_{t_0}^{T} \\
& - \int_{t_0}^{T} \left\{ (\lambda_*^T A)' D \delta x + H_x \delta x + H_u \delta u \right\} dt + o(\|\delta u\|_\infty).
\end{aligned} \tag{2.38}$$

When taking the limit $\epsilon \to 0$ and using the property (2.29) we obtain

$$\begin{aligned}
\delta J(u_*) \delta u \quad = \quad & \left(x_*^T(T) V + \lambda_*^T(T) A(T) D(T) \right) \delta x(T) \\
& - \int_{t_0}^{T} H_u \delta u \, dt - \int_{t_0}^{T} \{ H_x + (\lambda_*^T A)' D \} \delta x \, dt,
\end{aligned} \tag{2.39}$$

and since λ_* is a solution of (2.35), we have

$$\delta J(u_*) \delta u = - \int_{t_0}^{T} H_u \delta u \, dt. \tag{2.40}$$

As δu was an arbitrary admissible variation, (2.40) is valid for every variation $\delta u \in C([t_0, T], \mathbb{R}^m)$. The Fundamental Variation Lemma (e.g., [He]) provides, together with (2.32), the result

$$H_u^T(x_*, \lambda_*, u_*, t) = -S^T x_* + C^T \lambda_* - K u_* = 0. \tag{2.41}$$

Thus, it is proved that (x_*, u_*, λ_*) solves the boundary value problem (2.28). □

Our intention is to prove such a theorem also for the optimization problem with a more general controlled DAE, just making weak assumptions on the regularity and the index.

The first problem is that the set of admissible variations $(\delta x, \delta u)$ of a given optimal pair (x_*, u_*) is much more complicated than in the nice index 1 case. We just know that the directional derivate of J in (x_*, u_*) exists for an admissible direction $(\delta x, \delta u)$. The second problem is that we do not know in general if the final value problem (2.35) for the adjoint DAE is solvable. If the controlled DAE is regular with index $\mu > 2$, we even do not know anything about the regularity or the index of the adjoint DAE. If there is a an admissible variation $(\delta x, \delta u)$ and we have a solution λ_* to (2.35), then the directional derivate of J in (x_*, u_*) for the direction $(\delta x, \delta u)$ can be expressed by

$$\delta J(x_*, u_*) \begin{pmatrix} \delta x \\ \delta u \end{pmatrix} = \int_{t_0}^{T} \left\langle S^T x_* - C^T \lambda_* + K u_*, \delta u \right\rangle \, dt \tag{2.42}$$

similar to the expression (2.40) in the index 1 case. To stress it once again, the nice property of the index 1 case is that this holds for every $\delta u \in C([t_0, T], \mathbb{R}^k)$.

For example, in [Me] a linear state feedback is used in order to transform the general DAE into a new DAE with index 1. Then one can apply Theorem 2.12. The disadvantage of this approach is that it does not offer an explicit, necessary condition for the original DAE. Here, in Section 5, our main idea is to consider an extended system in order to obtain a larger DAE with index 1. The advantage is that the original DAE is preserved explicitly as a part of the extended system.

3. The DAE in extended Hessenberg-Form. Again we consider the DAE (2.1) with properly stated leading term. In [Mä1] the transformation of this DAE with nonsingular matrix functions $M, N \in C([t_0, T], \mathbb{R}^{m \times m})$ in the form

$$\tilde{A} := MA, \quad \tilde{D} := DN, \quad \tilde{B} := MBN, \quad \tilde{C} := MC \tag{3.1}$$

and the refactorization of the leading term with a matrix function $H \in C^1([t_0, T]\mathbb{R}^{r \times n})$ in the form

$$\tilde{A} := AH, \quad \tilde{D} := H^- D, \quad \tilde{B} := B + ADD^-(RH)' H^- D \tag{3.2}$$

is described. Here H^- is a generalized inverse of H such that $RHH^-R = R$. The generalized inverse D^- of D is chosen such that $D^-D = P_0$ and $DD^- = R$ (cf. [Mä1]).
The matrix M causes a scaling of the DAE while the matrix N means a transformation of the variable x in the form $\tilde{x} = N^{-1}x$. The matrix H leads to another factorization of the leading term that is also properly stated.
In [Mä1] it is shown that such a transformation in case of a regular DAE with index μ leads again to a regular DAE with index μ. In [Ba] a regular DAE with index 2 is transformed into so-called extended Hessenberg form by this transformation.

Looking at the transformation in [Ba] one can see that just with the assumption that G_1 has constant rank on $[t_0, T]$ the transformation to extended Hessenberg form is possible:

Theorem 3.1. *(DAE in extended Hessenberg form)*
For the DAE (2.1) with properly stated leading term we assume that G_1 has constant rank on the interval $[t_0, T]$. Then, by transformation and refactorization of the leading term, the DAE can be transformed into the extended Hessenberg form

$$\left\{ \begin{array}{rcl} \dot{x}_1 + B_{11}x_1 + B_{13}x_3 & = & C_1 u \\ x_2 & = & C_2 u \\ B_{31}x_1 & = & C_3 u. \end{array} \right. \tag{3.3}$$

Again we denote the transformed variable by x and consider $x_i \in \mathbb{R}^{m_i}$ for $i = 1, 2, 3$ with $m = m_1 + m_2 + m_3$. The matrices B_{11}, B_{13}, B_{31}, C_1, C_2, C_3 are assumed to have suitable dimensions.

For a regular DAE with index 1 a transformation into the form (3.3) would lead to the special case $m_3 = 0$. In the special case $m_2 = 0$ the DAE (3.3) is well known as a DAE in Hessenberg form. For a DAE in Hessenberg form it is also well known, that if the DAE (2.1) and, consequently, the DAE (3.3) are regular with index 2, the matrix $B_{31}B_{13} \in \mathbb{R}^{m_3 \times m_3}$ is nonsingular. Let us stress once again that we do not want to make this assumption here. We just consider DAEs that can be transformed into the form (3.3).
For example, in [ABUY] a very similar special form of the DAE which is basically the same as considered here, is used to analyze the DAE. Since in [ABUY] a more special class of DAEs is considered, also the transformation into the special form is a little bit more special than here.

In the tripartite form of the extended Hessenberg form the matrices V, W and S in the quadratic cost function are given by

$$V = \begin{pmatrix} V_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad W = \begin{pmatrix} W_{11} & W_{12} & W_{13} \\ W_{12}^T & W_{22} & W_{23} \\ W_{13}^T & W_{23}^T & W_{33} \end{pmatrix}, \quad S = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}. \tag{3.4}$$

The matrix K remains unchanged because it just valuates the non-transformed control function. The matrix V has its special structure with two zero rows because of property (2.12).

The corresponding adjoint equation

$$\begin{cases} \dot{\lambda}_1 - B_{11}^T\lambda_1 - B_{31}^T\lambda_3 &= W_{11}x_1 + W_{12}x_2 + W_{13}x_3 + S_1u \\ -\lambda_2 &= W_{12}^T x_1 + W_{22}x_2 + W_{23}x_3 + S_2u \\ -B_{13}^T\lambda_1 &= W_{13}^T x_1 + W_{23}^T x_2 + W_{33}x_3 + S_3u \end{cases} \tag{3.5}$$

is also of extended Hessenberg form and together with the condition

$$\lambda_1(T) = -V_{11}x_1(T) \tag{3.6}$$

it represents a final value problem that has to be solved.

4. The index of the DAE in the boundary value problem. As discussed in Section 2 under suitable conditions the sufficient condition from Theorem 2.9 provides the possibility to compute an optimal control just by solving the boundary value problem (2.15), which contains a large DAE with the variable (x, λ, u). So it is natural to ask for the index of the large DAE in (2.15). From [BaKuMä] we have a condition that is sufficient and necessary for the DAE in (2.15) to be regular with index 1:

THEOREM 4.1. *The DAE of the boundary value problem (2.15) is regular with index 1 if and only if the following two conditions are valid:*

$$im\,(G_1(t), C(t)) = \mathbb{R}^m \tag{4.1}$$

and

$$im \begin{pmatrix} G_0^T(t) - B^T(t)Q_{*0}(t) & W(t)Q_0(t) & S(t) \\ -C^T(t)Q_{*0}(t) & S^T(t)Q_0(t) & K(t) \end{pmatrix} = \mathbb{R}^m \times \mathbb{R}^k \tag{4.2}$$

for every $t \in [t_0, T]$.

In case the DAE (2.1) is in extended Hessenberg form (3.3), the matrices from (4.1) and (4.2) are of the form

$$(G_1, C) = \begin{pmatrix} I & 0 & B_{13} & C_1 \\ 0 & I & 0 & C_2 \\ 0 & 0 & 0 & C_3 \end{pmatrix} \tag{4.3}$$

$$\begin{pmatrix} G_0^T - B^T Q_{*0} & WQ_0 & S \\ -C^T Q_{*0} & S^T Q_0 & K \end{pmatrix} = \begin{pmatrix} I & 0 & -B_{31}^T & 0 & W_{12} & W_{13} & S_1 \\ 0 & -I & 0 & 0 & W_{22} & W_{23} & S_2 \\ 0 & 0 & 0 & 0 & W_{23}^T & W_{33} & S_3 \\ 0 & -C_2^T & -C_3^T & 0 & S_2^T & S_3^T & K \end{pmatrix}. \tag{4.4}$$

A transformation of the controlled DAE and a refactorization of the leading term also mean a transformation and refactorization of the leading term for the large DAE of the boundary value problem (cf. [Ba]). Here, again, the regularity and the index of the DAE are invariant. This yields the property

LEMMA 4.2. *The conditions (4.1) and (4.2) are invariant under transformation of the DAE resp. a refactorization of the leading term.*

The conditions (4.3) and (4.4) show that (4.1) and (4.2) are valid for the DAE in extended Hessenberg form if and only if

$$\begin{cases} \operatorname{Rang} C_3 = \operatorname{Rang} \begin{pmatrix} W_{23}^T & W_{33} & S_3 \end{pmatrix} = m_3 \\ \operatorname{Rang} \begin{pmatrix} C_2^T & C_3^T & S_2^T & S_3^T & K \end{pmatrix} = k \end{cases}. \tag{4.5}$$

In particular the case

$$k \geq m_3, \quad \operatorname{Rang} C_3 = \operatorname{Rang} W_{33} = m_3, \quad K \text{ nonsingular}, \tag{4.6}$$

is sufficient for (4.5)

Condition (4.2) is never fulfilled for the special cost function with (2.27). This can be seen in the extended Hessenberg form where $W_{23} = 0$, $W_{33} = 0$ und $S_3 = 0$ is valid for the special cost function.

5. The extended system. We consider a new DAE that is generated by extending the controlled DAE and we investigate the case when this new DAE is regular with index 1.

The extended system

We consider the idea to understand the control u as a variable and to combine the state and the control function to the new variable

$$z = \begin{pmatrix} x \\ u \end{pmatrix} \in \mathbb{R}^{m+k}. \tag{5.1}$$

For a matrix function $F \in C([t_0,T],\mathbb{R}^{k\times m})$ and functions $v \in C([t_0,T],\mathbb{R}^k)$ we consider the system

$$\left\{\begin{array}{rcl} A(Dx)' + Bx - Cu & = & 0 \\ Fx + u & = & v. \end{array}\right. \tag{5.2}$$

For the new variable z and control functions v the system (5.2) is again a controlled DAE with

$$\hat{A} := \begin{pmatrix} A \\ 0 \end{pmatrix}, \quad \hat{D} := \begin{pmatrix} D & 0 \end{pmatrix}, \quad \hat{B} := \begin{pmatrix} B & -C \\ F & I \end{pmatrix}, \quad \hat{C} := \begin{pmatrix} 0 \\ I \end{pmatrix}. \tag{5.3}$$

Thus, the DAE (5.2) has also a properly stated leading term.

<u>Index 1 for the extended system</u>

The question arises if it is possible to choose the matrix function F such that the DAE (5.2) is regular with index 1.

In this case we will consider the DAE (5.2) together with the initial value condition

$$\hat{A}(t_0)\hat{D}(t_0)(z(t_0) - z^0) = 0, \quad z^0 := \begin{pmatrix} \tilde{x}^0 \\ 0 \end{pmatrix}, \quad \tilde{x}^0 \in \mathbb{R}^m, \tag{5.4}$$

which offers us the nice Existence Theorem 2.2 for the solution of the initial value problem (5.2), (5.4).

For the matrix sequence belonging to (5.2) we have

$$\hat{G}_0 = \begin{pmatrix} AD & 0 \\ 0 & 0 \end{pmatrix}, \quad \hat{Q}_0 = \begin{pmatrix} Q_0 & 0 \\ 0 & I \end{pmatrix} \tag{5.5}$$

so that

$$\hat{G}_1 = \begin{pmatrix} G_1 & -C \\ FQ_0 & I \end{pmatrix}. \tag{5.6}$$

According to Definition 2.1 the DAE (5.2) is regular with index 1 if and only if the matrix $\hat{G}_1$ is nonsingular. We have

$$\begin{pmatrix} G_1 & -C \\ FQ_0 & I \end{pmatrix} \text{nonsingular} \iff AD + (B + CF)Q_0 \,\text{nonsingular}. \tag{5.7}$$

Thus, we have

Lemma 5.1. *The DAE (5.2) is regular with index 1 if and only if the matrix function $AD + (B + CF)Q_0$ is nonsingular on the interval $[t_0, T]$.*

Remark 5.2. *The condition from Lemma 5.1 is equal to the index-1-condition for the DAE*

$$A(Dx)' + (B + CF)x = Cv. \tag{5.8}$$

In [Me] the time invariant case is considered and the matrix F is chosen such that (5.8) has index 1. To this transformed DAE resp. to the transformed optimization problem we can apply Theorem 2.12. However, this approach does not provide an explicit necessary optimality condition for the old optimization problem. Together with the DAE the adjoint DAE is also transformed so that we do not obtain any information about the solvability of the original adjoint DAE. In system (5.2) the DAE (2.1) is preserved in its old form. We will see that this approach leads to information about the solvability of the adjoint equation.
The idea to consider the control as a variable in system (5.2) comes from the idea of the Behaviour Approach (e.g., [Wi]).

Now it is the question whether it is possible to choose the matrix function F such that (5.7) is fulfilled. Obviously, the condition

$$\operatorname{im}(G_1(t), C(t)) = \mathbb{R}^m \quad \text{for every} \quad t \in [t_0, T] \tag{5.9}$$

is necessary for (5.7). We want to check if (5.9) is also sufficient for (5.7).

In order to do this we consider the DAE (2.1) in extended Hessenberg form. Therefore according to Theorem 3.1, we must assume that G_1 has constant rank on the interval $[t_0, T]$.
For the DAE in extended Hessenberg form the matrix function F is of the form

$$F = \begin{pmatrix} F_1 & F_2 & F_3 \end{pmatrix} \tag{5.10}$$

and we have

$$\hat{G}_1 = \begin{pmatrix} I & 0 & B_{13} & -C_1 \\ 0 & I & 0 & -C_2 \\ 0 & 0 & 0 & -C_3 \\ 0 & F_2 & F_3 & I \end{pmatrix}. \tag{5.11}$$

Here we obtain the simple condition

$$\hat{G}_1 \quad \text{nonsingular} \quad \Longleftrightarrow \quad \begin{pmatrix} I & 0 & -C_2 \\ 0 & 0 & -C_3 \\ F_2 & F_3 & I \end{pmatrix} \quad \text{nonsingular.} \tag{5.12}$$

It is necessary for (5.12) that

$$\operatorname{Rang} C_3 = m_3 \tag{5.13}$$

and, therefore, $k \geq m_3$ has to hold, i.e., the dimension of the control has to be as large as the dimension of the solution component x_3 at least.

However, condition (5.13) is also sufficient for choosing F such that (5.12) is fulfilled. Choosing F as

$$F = (F_1, F_2, F_3) = (0, 0, C_3^T), \tag{5.14}$$

then $\hat{G}_1$ is nonsingular, as the matrix $C_3 C_3^T$ is nonsingular under the condition (5.13).

In order to obtain the desired F for the original DAE we have to invert the transformation into the extended Hessenberg form. For this, we set

$$\tilde{F} = FN. \tag{5.15}$$

Then the index-1-property of the DAE (5.2) is invariant with respect to the transformation and the refactorization of the leading term of the DAE (2.1). We have

$$\begin{pmatrix} \tilde{G}_1 & -\tilde{C} \\ \tilde{F}\tilde{Q}_0 & I \end{pmatrix} = \begin{pmatrix} M & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} G_1 & -C \\ FQ_0 & I \end{pmatrix} \begin{pmatrix} N & 0 \\ 0 & I \end{pmatrix}. \tag{5.16}$$

Finally, we obtain the following lemma, which provides index 1 for the extended system:

LEMMA 5.3. *For the DAE (2.1) we assume that G_1 has constant rank on the interval $[t_0, T]$. Then, by the choice of F, it is possible to obtain index 1 for the extended system (5.2) if and only if*

$$im\,(G_1(t), C(t)) = \mathbb{R}^m \tag{5.17}$$

is valid for every $t \in [t_0, T]$.

REMARK 5.4. *Condition (5.17) is equal to condition (4.1), which is necessary for the DAE of the boundary value problem to be regular with index 1.*

The extended system offers admissible pairs

Henceforth, we assume that, by the choice of F, the DAE (5.2) is regular with index 1. Thus we assume that G_1 has constant rank on the interval $[t_0, T]$.

We consider the set

$$\mathcal{A}(x^0) := \left\{ x \in \mathbb{R}^m | D(t_0)P[x - x^0] = 0 \right\} \tag{5.18}$$

of the values $\mathbb{R}^m$ that match the value x^0 in the $D(t_0)P$-component. This means, a function $x(t)$ satisfies the initial value condition (2.2) if and only if $x(t_0) \in \mathcal{A}(x^0)$.

Now we have

LEMMA 5.5. *Let $\tilde{x}^0 \in \mathcal{A}(x^0)$ and $v \in C([t_0, T], \mathbb{R}^k)$. Then the unique solution (x, u) of (5.2), (5.4) is an admissible pair for the initial value problem (2.1), (2.2).*

Proof. Since (x, u) is a solution of (5.2), (5.4), we have $A(Dx) + Bx - Cu = 0$ and $D(t_0)P[x(t_0) - x^0] = 0$ with $x \in C^1_D([t_0, T], \mathbb{R}^m)$ and $u \in C([t_0, T], \mathbb{R}^k)$, i.e., (x, u) is an admissible pair. ∎

For the solution x of (2.1), (2.2) from Lemma 5.5 one has to notice that

$$D(t_0)x(t_0) = D(t_0)Px(t_0) + D(t_0)Qx(t_0) \tag{5.19}$$

and here only the $D(t_0)P$-component of $x(t_0)$ is given by the initial value condition (2.2), whereas the $D(t_0)Q$-component of $x(t_0)$ is a value that has to be optimized in the optimization problem (2.13).

In addition to Lemma 5.5 we also have

LEMMA 5.6. *Let* (x,u) *be an admissible pair for the initial value problem (2.1), (2.2). Then* (x,u) *is the unique solution of the initial value problem (5.2), (5.4) with* $\tilde{x}^0 := x(t_0)$ *and* $v := Fx + u$.

Proof. Since (x,u) is an admissible pair, we have $A(Dx)' + Bx = Cu$ with $x \in C^1_D([t_0,T],\mathbb{R}^m)$ and $u \in C([t_0,T],\mathbb{R}^k)$, i.e., $(x,u) \in C^1_{\hat{D}}([t_0,T],\mathbb{R}^k)$. Thus, by the choice of $\tilde{x}^0$ and v we obtain (x,u) as solution of (5.2), (5.4). □

As a result of Lemma 5.5 and 5.6 we derive, theoretically, the whole set of admissible pairs (x,u) as solutions of (5.2), (5.4) if $v \in C([t_0,T],\mathbb{R}^k)$ and $\tilde{x}^0 \in \mathcal{A}(x^0)$ pass through all possible values.

The necessary optimality condition for (2.13)

We want to prove a necessary optimality condition for the optimization problem (2.13). For this purpose we start with an optimal pair for (2.13) and show that the optimal pair provides an optimal control to an optimization problem for the system (5.2). As (5.2) is regular with index 1, the well-known necessary optimality condition from Theorem 2.12 is valid, and from this we want to derive the desired necessary optimality condition for (2.13).

Hence, let us assume that (x_*,u_*) is an optimal pair for the optimization problem (2.13). We want to define a linear-quadratic optimization problem for the DAE (5.2). For this purpose we choose the value

$$\tilde{x}^0 := x_*(t_0) \in \mathcal{A}(x^0) \tag{5.20}$$

of the optimal trajectory x_* in t_0 for the initial value condition (5.4) and we obtain the initial value condition

$$\hat{D}(t_0)z(t_0) = D(t_0)x(t_0) = D(t_0)x_*(t_0). \tag{5.21}$$

Let $v \in C([t_0,T],\mathbb{R}^k)$ be a control function for the DAE (5.2) and (x,u) the corresponding unique admissible pair from Lemma 5.5. Then, for the cost function (2.11) of the optimization problem (2.13), we have

$$\begin{aligned} J(x,u) &= \tfrac{1}{2}\begin{pmatrix} x(T) \\ u(T) \end{pmatrix}^T \begin{pmatrix} V & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x(T) \\ u(T) \end{pmatrix} \\ &\quad + \tfrac{1}{2}\int_{t_0}^T \begin{pmatrix} x \\ u \end{pmatrix}^T \begin{pmatrix} W & S \\ S^T & K \end{pmatrix}\begin{pmatrix} x \\ u \end{pmatrix} dt, \end{aligned} \tag{5.22}$$

which yields a quadratic cost function $\hat{J}$ for the initial value problem (5.2), (5.4) with

$$\hat{V} = \begin{pmatrix} V & 0 \\ 0 & 0 \end{pmatrix}, \quad \hat{W} = \begin{pmatrix} W & S \\ S^T & K \end{pmatrix}, \quad \hat{S} = 0, \quad \hat{K} = 0 \tag{5.23}$$

such that

$$\hat{J}(v) = J(x,u). \tag{5.24}$$

Thus, we obtain the optimization problem

$$\hat{J}(v) \to \mathrm{Min} \quad , \quad v \in C([t_0,T],\mathbb{R}^k), \tag{5.25}$$

and we can prove the following theorem

THEOREM 5.7. *(optimal pair (x_*,u_*) provides v_*)*
Assume that (x_*,u_*) is an optimal pair for the optimization problem (2.13). Then

$$v_* := Fx_* + u_* \tag{5.26}$$

is an optimal control for the optimization problem (5.25) and $z_* = (x_*,u_*)$ is the corresponding optimal trajectory.

Proof. $z_* = (x_*,u_*)$ is a solution for (5.2), (5.4) with the control function v_*. We consider a variation $\delta v \in C([t_0,T],\mathbb{R}^k)$ of v_*. Let $(x_* + \delta x, u_* + \delta u)$ be the solution of (5.2) with the control function $v_* + \delta v$ and $D(t_0)\delta x(t_0) = 0$. Then $(\delta u, \delta x)$ is an admissible variation of (x_*,u_*). Thus, for a variation $\delta v \in C([t_0,T],\mathbb{R}^k)$ of v_*, we obtain

$$\hat{J}(v_* + \delta v) = J(x_* + \delta x, u_* + \delta u) \geq J(x_*,u_*) = \hat{J}(v_*) \tag{5.27}$$

and, hence, v_* is an optimal control for (5.25) and $z_* = (x_*,u_*)$ is the corresponding optimal trajectory. ∎

Now consider u_*, x_* and v_* as in Theorem 5.7. Then the adjoint DAE to the optimization problem (5.25)

$$\hat{D}^T(\hat{A}\mu)' - \hat{B}^T\mu = \hat{W}z_* + \hat{S}v_* \tag{5.28}$$

resp.

$$\left\{ \begin{array}{rcl} D^T(A^T\mu_1)' - B^T\mu_1 - F^T\mu_2 & = & Wx_* + Su_* \\ C^T\mu_1 - \mu_2 & = & S^T x_* + Ku_* \end{array} \right. \tag{5.29}$$

is regular with index 1 and has a continuous right-hand side. Thus, according to Theorem 2.2, equation (5.29) together with the final value condition

$$D^T(T)A^T(T)\mu_1(T) = -Vx_*(T) \tag{5.30}$$

has a unique solution

$$\mu_* \in C^1_{\hat{A}^T}([t_0,T],\mathbb{R}^{m+k}). \tag{5.31}$$

As (5.25) is an optimization problem for the index-1-DAE (5.2), the necessary optimality condition from Theorem 2.12 is fulfilled here and we have

$$\hat{S}^T z_* - \hat{C}^T\mu_* + \hat{K}v_* = 0 \quad \Leftrightarrow \quad \hat{C}^T\mu_* = 0 \quad \Leftrightarrow \quad \mu_{*2} = 0. \tag{5.32}$$

According to the first equation in (5.29) we obtain

$$\lambda_* := \mu_{*1} \in C^1_{A^T}([t_0,T],\mathbb{R}^m) \tag{5.33}$$

as a solution of the final value problem

$$\begin{cases} D^T(A^T\lambda)' - B^T\lambda = Wx + Su \\ D^T(T)A^T(T)\lambda(T) = -Vx_*(T) \end{cases} \tag{5.34}$$

and, having this, we obtain according to the second equation in (5.29) that

$$S^T x_* - C^T\lambda_* + Ku_* = 0. \tag{5.35}$$

Finally, we have proved the following Theorem:

THEOREM 5.8. *(Necessary optimality condition)*
For the optimization problem (2.13), assume that

1. G_1 *has constant rank on the interval* $[t_0, T]$*;*
2. (G_1, C) *has full rank on the interval* $[t_0, T]$.

Assume that (x_*, u_*) *is an optimal pair, i.e.,* $u_* \in C([t_0,T],\mathbb{R}^k)$ *is an optimal control and* $x_* \in C^1_D([t_0,T],\mathbb{R}^m)$ *is a corresponding optimal trajectory. Then there exists a function*

$$\lambda_* \in C^1_{A^T}([t_0,T],\mathbb{R}^m)$$

such that (x_*, λ_*, u_*) *is a solution of the boundary value problem*

$$\left\{\begin{array}{rcl} A(Dx)' &=& -Bx + Cu \\ D^T(A^T\lambda)' &=& Wx + B^T\lambda + Su \\ 0 &=& S^Tx - C^T\lambda + Ku \\ && \\ D(t_0)P\left[x(t_0) - x^0\right] &=& 0 \\ D^T(T)A^T(T)\lambda(T) &=& -Vx(T). \end{array}\right. \tag{5.36}$$

6. Conclusion. Theorem 5.8 offers the guaranty that, under corresponding assumptions, an existing optimal control can be computed by computing the corresponding solution of the boundary value problem (5.36). Note that there might be a problem if (5.36) has more than one solution.
Vice versa, it is important to notice once again that the sufficient condition is only valid in case $D(t_0)P = D(t_0)$ in general. This means that if $D(t_0)P \neq D(t_0)$, a solution of (5.36) does not need to offer an optimal control.
Notice once again that according to Theorem 4.1 the assumption that (G_1, C) has full rank is fulfilled especially in the case if the DAE in (5.36) is regular with index 1. Theorem 4.1 also shows that, vice versa, we need additional assumptions on the cost function (more precisely, on the matrices W and S) in order to assure index 1 for the DAE in (5.36).

Finally, the situation is fine for an optimization problem that fulfils the assumptions of Theorem 4.1 and uses the projector $P = P_0(t_0)$ in the initial value condition. In this case we have an optimality condition that is both necessary and sufficient and the DAE in (5.36) that has to be solved is regular with index 1. This is a suitable situation for a numerical method to compute an optimal control. Note once again that, for this situation, no explicit assumption concerning the index and the regularity of the controlled DAE are necessary except for the mild condition that the matrix G_1 has constant rank.

REFERENCES

[ABUY] A. Abramov, K.Balla, V.I. Ulynova, L.F. Yukhno: On a nonlinear selfadjoint eigenvalue problem for a certain class of differential algebraic equations, Comp. Math. Math. Phys., 43(3): 391-402, 2003. Translated from Zh. Vychisl. Mat. Mat. Fiz. 43(3): 410-421,2003

[Ba] A. Backes: Optimale Steuerung der linearen DAE im Fall Index 2, Humboldt University Berlin, Institute of Mathematics, Preprint 2003-04

[BaKuMä] K. Balla, G. Kurina, R. März: Index criteria for differential algebraic equations arising from linear-quadratic optimal control problems, Humboldt University Berlin, Institute of Mathematics, Preprint 2003-14

[BaMä] K. Balla und R. März: A Unified approach to linear differential algebraic equations and their adjoint equations. Zeitschrift für Analysis und ihre Anwendungen 21(2002)3, 783-802

[BeLa] D. Bender, A.Laub: The linear-quadratic optimal regulator for descriptor systems, IEEE Transactions on Automatic Control, Vol.32, No.8, 672-688, 1987

[Cobb] D. Cobb: Descriptor variable systems and optimal state regulation, IEEE Transactions on Automatic Control, Vol.28, No.5, 1983

[He] Hestenes, Magnus R.: Calculus of variations and optimal control theory, New York, 1966

[Jo] E. Jonckheere: Variational Calculus for descriptor problems, IEEE Transactions on Automatic Control, Vol.33, No.5, 1988

[KuMe] P. Kunkel, V. Mehrmann: The linear-quadratic optimal control problem for linear descriptor systems with variable coefficients, Math. Control of Signals and Systems, Vol.10, No.3, 1997

[KuMä] G.A. Kurina, R. März: On linear-quadratic optimal control problems for time-varying descriptor systems, SIAM J. Control and Optimization, to appear

[Mä1] R. März: The index of linear differential algebraic equations with properly stated leading terms. Results in Mathmatics 42(2002), 308-338

[Mä2] R. März: Adjoint equations of differential-algebraic systems and optimal control problems, Pro. Belarussian National Academy of Sciences, Institute of Mathematics, Minsk, Vol. 7, 88-97, 2001

[Mä3] R. März: Differential algebraic systems anew. Applied Numerial Mathmatics 42(2002), 315-335

[Mä4] R. März: Solvability of linear differential algebraic equations with properly stated leading terms. Humboldt University Berlin, Institute of Mathematics, Preprint 2002-12

[Me] V. Mehrmann: The autonomous linear-quadratic control problem, Lecture Notes in Control and Information Science, 163, Springer-Verlag, 1991

[Wi] J.C. Willems: Paradigms and Puzzles in the Theoty of Dynamical Systems, IEEE Transactions on Automatic Control, Vol. 36, No.3, 259-294, 1991